What’s different about the State Summative Assessment (SSA) for science?

Phenomenon-based item clusters

The field test of the SSA will utilize a phenomenon-based approach. Rather than being a collection of individual and unrelated items, questions will be grouped in **clusters** in which every question will be related to the others by asking students to use their understanding of science and/or engineering design to consider a phenomenon or solve an engineering design problem. A **phenomenon** is anything that can be observed and that students can attempt to understand or explain.

Each cluster of items will attempt to gather information about two or more standards that have been **bundled** together by the item writer. The phenomenon chosen by the item writers is one that allows the bundled standards to be assessed in a logical way.

The assessment clusters use narrative text to lead students through the progression of questions. This narrative text is referred to as the **storyline** and its purpose is to set up the context of the phenomenon and to link the items together in a way that makes sense to the student.

For example, the phenomenon of stream erosion could be used to create a cluster of items. The two standards bundled together to measure by this cluster could be an ESS (earth science) standard and one from ETS (engineering design) to create a cluster that asks students to both explain the reason a stream becomes muddy and to evaluate possible solutions to prevent this from happening. The storyline for this cluster might involve two friends noticing the creek near their house becoming muddy after a rain storm and attempting to explain why, then brainstorming ways to reduce the amount of sediment that washes into the stream.

One immediate difference students will notice when opening the field test booklets is this storyline. Rather than starting with Question 1 on the first page, each cluster begins with varying amounts of text that establish the phenomenon and/or provide needed information for students to answer the questions that follow. For the hypothetical example above the storyline might begin:

> *Two friends, Marcus and Tammy are walking down their street when they notice the stream that flows beside it is much higher than normal, and the water is dark brown and muddy.*
> *“Wow, look at the stream,” said Marcus. “I’ve never seen it that muddy before.”*
> *“Yeah,” said Tammy. “It was really clear last week. I wonder what changed to make it so muddy.”*
> *“Let’s see if we can find some information to help us figure it out.”*

This text might then be followed with a table of rainfall data, and then the first question requiring students to use the data to provide an answer.

Multi-dimensional items

The other difference as compared to previous science assessments students are familiar with is the multi-dimensional approach to item development. Item writers were asked to develop two or three dimensional items, even for multiple choice items. This means the items are designed to measure students ability to use the Science and Engineering Practices (SEP) and Crosscutting Concepts (CCC) as well as the Disciplinary Core Ideas (DCI) of the standards. Most science assessments in the past were heavily weighted toward what is traditionally thought of as science "content" and few items attempted to measure the practice of science. On the field test students will be asked to identify patterns, construct arguments, link cause and effect, etc. *using* the DCI as the context for applying those SEP and CCC.

For example:

Which of these most likely caused the stream to become muddy?

a) Rain washes soil into the stream

as compared to

Tammy claimed that the rainfall data explains why the stream becomes muddy. What pattern in the rainfall data supports her claim?

a) The creek was always muddy 2-3 days after it rained upstream

In the second instance there is a focus on the student being able to identify a pattern in the data. Patterns are one of the Crosscutting Concepts, so the second question is an example of how a question can assess both a CCC as well as the DCI.

Specifics of test construction

In general, each cluster will consist of six (6) multiple choice questions and two (2) open response. This is not an absolute, as there are some clusters with more or less of both question type as determined by the standards that were bundled to create the clusters.

Some multiple choice questions require students to select more than one correct answer. Unless the item specifies otherwise, there is a single correct answer. The number of correct response will be identified in every question as follows:

For single answer questions the quantifier *one* will be included in the question:

(7) Which one location would be the **best** place to build a dam?

For multiple answer questions the number of correct responses will be specified in a separate line immediately preceding the answer choices:

(7) Which locations for a dam meet the design criteria?

Select the TWO **best** answers.

Multiple choice questions are not limited to a fixed number of answer choices. Although the large majority contain four (4) there are some that contain more, especially in the cases where students are asked to select more than one answer.

Constructed response questions on the field test ask for a greater variety of response types than in previous assessments. Students may be asked to draw and annotate a model, create a flow chart, create and explain a graph or some other type of response beyond a traditional written answer. In some cases, constructed response questions may incorporate multiple elements such as several related short answer questions combined within the same answer space. For example:

> Part A. Where would the water collect if Gate A were placed in the following positions: *(your answer choices could include upper pool, lower pool, river, reservoir, or unchanged)*
>
> 1. Position X ____
> 2. Position Y ____
> 3. Position Z ____
>
> Part B. Where would the water collect if Gate B were placed in the following positions: *(your answer choices could include upper pool, lower pool, river, reservoir, or unchanged)*
>
> 4. Position M____
> 5. Position Q ____
> 6. Position R ____
>
> Part C. Based on these results, what predictions can you make about the effect of Gate C on the system?

Item vocabulary

At all grade levels, the terminology of the questions reflects the language of the standards. This includes the wording of the SEP and CCC as well as any content-specific language contained in the Kentucky Academic Standards (KAS) for Science.

For instance, the first grade standard 1-LS1-1 contains the word “mimicking” and therefore it might be possible for this word to appear in a question. Fourth grade standard 4-PS4-1 contains “amplitude” and “wavelength,” so students might be expected to understand how both terms relate to the properties of a wave.

Common terminology used in the SEP and CCC might also be expected to appear in questions, such as asking students to “construct an explanation” or to “generate a solution.” Since the Engineering Design (ETS) standards will be included in the field test, terms such as “criteria for success” and “constraints” may be included in questions. For instance, a student might be asked:

> *Which one of these criteria for success would* **most likely** *be met by Joe’s proposed solution?*

Teachers should **let the KAS for Science be their guide for understanding the potential vocabulary of the assessment**, but are cautioned not to develop definition lists because the assessment will not require students to define science vocabulary terms. Forcing a student to memorize a formal and elaborate definition of “constraints” is an unnecessary use of valuable instructional time. A better use would be to simply make sure students understand that a constraint is a limit, and to provide relevant examples. For instance the school lunch is constrained by a 30 minute lunch period, and by the daily menu or the contents of their lunchbox.